

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 17:20:04 ; Search time 34 Seconds  
(without alignments)  
30.359 Million cell updates/sec

Title: US-09-807-980-8  
Perfect score: 24  
Sequence: 1 GGGF 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query					
No.	Score	Match	Length	DB	ID		Description
-----							

1	24	100.0	19	2	O05601	O05601 pseudomonas
2	24	100.0	25	10	Q42086	Q42086 arabidopsis
3	24	100.0	26	6	Q9TRE4	Q9tre4 bos taurus
4	24	100.0	29	16	Q9KAV1	Q9kav1 bacillus ha
5	24	100.0	30	2	Q9R7L0	Q9r7l0 escherichia
6	24	100.0	35	2	Q47874	Q47874 eubacterium
7	24	100.0	38	12	O89832	O89832 budgerigar
8	24	100.0	38	12	O89905	O89905 budgerigar
9	24	100.0	38	12	O89834	O89834 budgerigar
10	24	100.0	38	12	O89826	O89826 budgerigar
11	24	100.0	38	12	O89836	O89836 budgerigar
12	24	100.0	38	12	O89827	O89827 budgerigar
13	24	100.0	38	12	O89823	O89823 budgerigar
14	24	100.0	38	12	O89830	O89830 budgerigar
15	24	100.0	38	12	O89825	O89825 budgerigar
16	24	100.0	38	12	O89829	O89829 budgerigar
17	24	100.0	38	12	O89833	O89833 budgerigar
18	24	100.0	38	12	O89831	O89831 budgerigar
19	24	100.0	38	12	O89824	O89824 budgerigar
20	24	100.0	38	12	O89822	O89822 budgerigar
21	24	100.0	38	12	O89835	O89835 budgerigar
22	24	100.0	42	10	O23471	O23471 arabidopsis
23	24	100.0	43	11	Q63281	Q63281 rattus norv
24	24	100.0	43	16	Q9KF30	Q9kf30 bacillus ha
25	24	100.0	44	16	Q8UDT2	Q8udt2 agrobacteri
26	24	100.0	46	12	Q8JL3	Q8jjl3 hepatitis e
27	24	100.0	46	12	O55688	O55688 hepatitis e
28	24	100.0	46	12	O92647	O92647 hepatitis e
29	24	100.0	47	10	Q9S857	Q9s857 glycine max
30	24	100.0	47	16	Q9K9F6	Q9k9f6 bacillus ha
31	24	100.0	48	11	Q60413	Q60413 cricetus cr
32	24	100.0	52	2	Q9ZA99	Q9za99 pseudomonas
33	24	100.0	54	2	Q8VQ62	Q8vq62 bacillus li
34	24	100.0	56	2	Q9RMH8	Q9rmh8 lactobacill
35	24	100.0	56	12	O39853	O39853 grapevine 1
36	24	100.0	57	5	Q9U3E7	Q9u3e7 caenorhabdi
37	24	100.0	57	6	P83095	P83095 bos taurus
38	24	100.0	57	12	Q8JL8	Q8jjl8 hepatitis e
39	24	100.0	60	6	Q9GMN4	Q9gm4 macaca fasc
40	24	100.0	60	10	Q9ZP98	Q9zp98 hordeum vul
41	24	100.0	61	2	Q8KK56	Q8kk56 proteus vul
42	24	100.0	61	16	Q8XQD3	Q8xqd3 ralstonia s
43	24	100.0	62	5	Q25250	Q25250 lucilia cup
44	24	100.0	64	5	Q25245	Q25245 lucilia cup
45	24	100.0	64	5	Q25251	Q25251 lucilia cup

# ALIGNMENTS

## RESULT 1

O05601

ID	O05601	PRELIMINARY;	PRT;	19 AA.
AC	O05601;			
DT	01-JUL-1997	(TrEMBLrel. 04, Created)		
DT	01-JUL-1997	(TrEMBLrel. 04, Last sequence update)		
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)		

DE Transposon Tn5041 DNA (Fragment).  
 OS Pseudomonas sp.  
 OC Bacteria; Proteobacteria.  
 OX NCBI\_TaxID=306;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KHP41; TRANSPOSON=Tn5041;  
 RX MEDLINE=97419493; PubMed=9274008;  
 RA Khloldii G.Y., Yurieva O.V., Gorlenko Z.M., Mindlin S.Z., Bass I.A.,  
 RA Lomovskaya O.L., Kopteva A.V., Nikiforov V.G.;  
 RT "Tn5041 : a chimeric mercury resistance transposon closely related to  
 RT the toluene degradative transposon Tn4651.";  
 RL Microbiology 143:2549-2556(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KHP41; TRANSPOSON=Tn5041;  
 RA Kholodii G.Y., Mindlin S.Z., Gorlenko Z.M., Bass I.A., Kalyaeva E.S.,  
 RA Nikiforov V.;  
 RT "Host-dependent transposition of Tn5041.";  
 RL Russ. J. Genet. 36:365-373(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KHP41; TRANSPOSON=Tn5041;  
 RA Kholodii G.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; X98999; CAA67457.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 19 19  
 SQ SEQUENCE 19 AA; 1797 MW; 2070A562EF85EDD9 CRC64;

Query Match 100.0%; Score 24; DB 2; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGF 4  
 ||||  
 Db 3 GGGF 6

Search completed: October 20, 2003, 17:24:31  
 Job time : 38 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 17:19:34 ; Search time 11 Seconds  
(without alignments)  
17.101 Million cell updates/sec

Title: US-09-807-980-8  
Perfect score: 24  
Sequence: 1 GGGF 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	24	100.0	31	1	FBRL_RAT P22509 rattus norv
2	24	100.0	48	1	DHSU_CHLLT P20161 chlorobium
3	24	100.0	50	1	RT11_BOVIN P82911 bos taurus
4	24	100.0	60	1	CD52_MACFA P32763 macaca fasc
5	24	100.0	62	1	YLAF_BACSU O07630 bacillus su
6	24	100.0	70	1	FLC1_ECOLI P11519 escherichia
7	24	100.0	70	1	FLC2_ECOLI P23587 escherichia
8	24	100.0	79	1	DCOR_PARBR Q92445 paracoccidi
9	24	100.0	82	1	YEAQ_ECOLI P76246 escherichia
10	24	100.0	84	1	TAG1_ECO27 P58767 escherichia
11	24	100.0	84	1	TAG1_ECOLI P76011 escherichia
12	24	100.0	85	1	RT12_LEITA Q34940 leishmania
13	24	100.0	87	1	RT12_TRYBO Q33569 trypanoplas
14	24	100.0	89	1	SLTB_BPH30 P08027 bacterioph
15	24	100.0	90	1	IM10_NEUCR Q9c0n3 neurospora
16	24	100.0	90	1	Y585_ARCFU O29670 archaeoglob
17	24	100.0	98	1	KRFT_LARNO P02451 larus novae

18	24	100.0	104	1	HOL3_HOLDI	Q25055	holotrichia
19	24	100.0	108	1	IM13_CAEEL	O45319	caenorhabdi
20	24	100.0	111	1	USPB_ECOLI	P37632	escherichia
21	24	100.0	113	1	SSIF_STRBI	Q9r645	streptomyce
22	24	100.0	113	1	YF65_SYNY3	P74596	synechocyst
23	24	100.0	114	1	YC83_PORPU	P51217	porphyra pu
24	24	100.0	116	1	Y105_ARATH	Q9xik3	arabidopsis
25	24	100.0	116	1	YI57_AQUAE	O67709	aquifex aeo
26	24	100.0	118	1	RLA3_ORYSA	P56724	oryza sativ
27	24	100.0	118	1	YE17_SYNY3	P72731	synechocyst
28	24	100.0	120	1	YUTM_BACSU	O32113	bacillus su
29	24	100.0	121	1	HV01_MOUSE	P01745	mus musculu
30	24	100.0	129	1	RS9_BACST	P07842	bacillus st
31	24	100.0	130	1	RS9_BACHD	Q9kgd4	bacillus ha
32	24	100.0	130	1	RS9_CLOAB	Q97el3	clostridium
33	24	100.0	130	1	RS9_STAAM	Q99s52	staphylococ
34	24	100.0	131	1	RK12_CHLVU	P56345	chlorella v
35	24	100.0	131	1	RS9_HALN1	Q9hqj2	halobacteri
36	24	100.0	132	1	RS9_MYCGE	P47656	mycoplasma
37	24	100.0	132	1	RS9_MYCPN	P75179	mycoplasma
38	24	100.0	133	1	RS9_UREPA	Q9ppr3	ureaplasma
39	24	100.0	135	1	RS9_ARCFU	O29136	archaeoglob
40	24	100.0	135	1	RS9_PYRAB	Q9v195	pyrococcus
41	24	100.0	135	1	RS9_PYRFU	Q8u0e7	pyrococcus
42	24	100.0	135	1	RS9_PYRHO	O59299	pyrococcus
43	24	100.0	141	1	CHLE_CANFA	P32750	canis famil
44	24	100.0	141	1	CHLE_MACMU	P32751	macaca mula
45	24	100.0	141	1	CHLE_PIG	P32752	sus scrofa

# ALIGNMENTS

## RESULT 1

### FBRL\_RAT

ID FBRL\_RAT STANDARD; PRT; 31 AA.  
AC P22509;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Fibrillarin (Nucleolar protein 1) (Fragment).  
GN FBL.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=86033920; PubMed=2414294;  
RA Lischwe M.A., Ochs R.L., Reddy R., Cook R.G., Yeoman L.C., Tan E.M.,  
RA Reichlin M., Busch H.;  
RT "Purification and partial characterization of a nucleolar scleroderma  
RT antigen (Mr = 34,000; pI, 8.5) rich in NG,NG-dimethylarginine.";  
RL J. Biol. Chem. 260:14304-14310(1985).  
RN [2]  
RP SEQUENCE OF 1-28, AND INTERACTION WITH NOLC1.  
RX MEDLINE=20143579; PubMed=10679015;

RA Yang Y., Isaac C., Wang C., Dragon F., Pogacic V., Meier U.T.;  
 RT "Conserved composition of mammalian box H/ACA and box C/D small  
 RT nucleolar ribonucleoprotein particles and their interaction with the  
 RT common factor Nopp140."  
 RL Mol. Biol. Cell 11:567-577(2000).  
 CC -!- FUNCTION: FIBRILLARIN IS A COMPONENT OF A NUCLEOLAR SMALL NUCLEAR  
 CC RIBONUCLEOPROTEIN PARTICLE THOUGHT TO PARTICIPATE IN THE FIRST  
 CC STEP IN PROCESSING PRERIBOSOMAL RNA. IT IS ASSOCIATED WITH THE  
 CC U3, U8 AND U13 SMALL NUCLEAR RNAS.  
 CC -!- SUBUNIT: Interacts with Nolc1/Nopp140.  
 CC -!- SUBCELLULAR LOCATION: Nuclear; fibrillar region of the nucleolus.  
 CC -!- SIMILARITY: BELONGS TO THE FIBRILLARIN FAMILY.  
 DR PIR; A23887; A23887.  
 DR InterPro; IPR000692; Fibrillarin.  
 DR PROSITE; PS00566; FIBRILLARIN; PARTIAL.  
 KW Nuclear protein; Methylation; Ribonucleoprotein; rRNA processing;  
 KW RNA-binding.  
 FT MOD\_RES 8 8 METHYLATION (DI-).  
 FT MOD\_RES 15 15 METHYLATION (DI-).  
 FT MOD\_RES 21 21 METHYLATION (DI-).  
 FT MOD\_RES 24 24 METHYLATION (DI-).  
 FT MOD\_RES 28 28 METHYLATION (DI-).  
 FT MOD\_RES 31 31 METHYLATION (DI-).  
 FT CONFLICT 2 2 K -> D (IN REF. 2).  
 FT NON\_TER 31 31  
 SQ SEQUENCE 31 AA; 2965 MW; 93EBCC102847A363 CRC64;

Query Match 100.0%; Score 24; DB 1; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGF 4  
 ||||  
 Db 9 GGGF 12

Search completed: October 20, 2003, 17:23:49  
 Job time : 13 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 17:22:04 ; Search time 15 Seconds  
(without alignments)  
25.645 Million cell updates/sec

Title: US-09-807-980-8  
Perfect score: 24  
Sequence: 1 GGGF 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8							
Result	Query						
No.	Score	Match	Length	DB	ID	Description	
1	24	100.0	10	2	E41946	T-cell receptor ga	
2	24	100.0	16	2	PH1453	T-cell receptor al	
3	24	100.0	29	2	A83923	hypothetical prote	
4	24	100.0	31	2	A23887	fibrillarin - rat	
5	24	100.0	31	2	A21919	tricarboxylate-bin	
6	24	100.0	35	2	S18264	(S)-6-hydroxynicot	
7	24	100.0	42	2	A71429	hypothetical prote	
8	24	100.0	43	2	F83732	hypothetical prote	
9	24	100.0	44	2	AG2826	hypothetical prote	
10	24	100.0	47	2	C83986	hypothetical prote	
11	24	100.0	57	2	T22546	hypothetical prote	
12	24	100.0	60	2	S27152	GPI-anchored epidi	
13	24	100.0	62	2	D69872	hypothetical prote	

14	24	100.0	66	2	H81804	hypothetical prote
15	24	100.0	66	2	B81061	hypothetical prote
16	24	100.0	67	2	B83792	small acid-soluble
17	24	100.0	70	2	T42189	fatty-acid-CoA lig
18	24	100.0	70	2	JS0433	F-plasmid maintena
19	24	100.0	73	2	JS0107	hypothetical 7.4K
20	24	100.0	74	2	T22077	hypothetical prote
21	24	100.0	74	2	T33087	hypothetical prote
22	24	100.0	76	2	T32666	hypothetical prote
23	24	100.0	78	2	E81116	probable lipoprote
24	24	100.0	82	2	C64940	transglycosylase-a
25	24	100.0	82	2	D85790	transglycosylase-a
26	24	100.0	82	2	AG0713	probable membrane
27	24	100.0	82	2	H90941	transglycosylase-a
28	24	100.0	84	2	H64865	transglycosylase-a
29	24	100.0	84	2	D85698	hypothetical prote
30	24	100.0	84	2	D90840	hypothetical prote
31	24	100.0	84	2	AF0722	probable membrane
32	24	100.0	84	2	S33999	EP84R protein - Af
33	24	100.0	84	2	AC3279	hypothetical prote
34	24	100.0	85	2	T32664	hypothetical prote
35	24	100.0	85	2	F64529	hypothetical prote
36	24	100.0	85	2	A42056	ribosomal protein
37	24	100.0	86	2	C71960	hypothetical prote
38	24	100.0	86	2	F86294	hypothetical prote
39	24	100.0	87	2	T26867	hypothetical prote
40	24	100.0	89	1	XVEBBD	Shigella toxin cha
41	24	100.0	89	1	JN0726	Shiga-like toxin I
42	24	100.0	89	1	XVBPH9	Shiga-like toxin c
43	24	100.0	89	2	B53887	Shiga-like toxin I
44	24	100.0	89	2	E91000	Shiga toxin I subu
45	24	100.0	89	2	G85845	Shiga toxin I subu

#### ALIGNMENTS

##### RESULT 1

E41946

T-cell receptor gamma chain (1a.9) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-May-1999

C;Accession: E41946

R;Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.

Mol. Cell. Biol. 11, 5902-5909, 1991

A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma genes in intestinal intraepithelial lymphocytes from murine athymic chimeras.

A;Reference number: A41946; MUID:92049316; PMID:1658619

A;Accession: E41946

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-10 <WHE>

C;Keywords: T-cell receptor

Query Match 100.0%; Score 24; DB 2; Length 10;



Best Local Similarity 100.0%; Pred. No. 77;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGF 4  
    ||||  
Db 7 GGGF 10

Search completed: October 20, 2003, 17:24:52  
Job time : 17 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 17:32:30 ; Search time 95 Seconds  
(without alignments)  
10.865 Million cell updates/sec

Title: US-09-807-980-8  
Perfect score: 24  
Sequence: 1 GGGF 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 452

Minimum DB seq length: 0  
Maximum DB seq length: 8

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertibrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	
No.	Score Match Length DB ID	Description
<hr/>		

1	12	50.0	7	11	Q63480	Q63480 rattus norv
2	12	50.0	7	11	O55184	O55184 rattus norv
3	12	50.0	8	4	Q15901	Q15901 homo sapien
4	12	50.0	8	11	Q9QVI5	Q9qvi5 rattus sp.
5	12	50.0	8	11	Q9R1U6	Q9rlu6 mus musculu
6	11	45.8	7	10	O49223	O49223 glycine max
7	11	45.8	8	6	Q9TT78	Q9tt78 canis famil
8	11	45.8	8	8	P92386	P92386 hordeum mar
9	10	41.7	8	5	P82688	P82688 periplaneta
10	9	37.5	8	2	Q9R5R2	Q9r5r2 shigella dy
11	9	37.5	8	7	Q95213	Q95213 oryctolagus
12	8	33.3	8	8	P93963	P93963 psathyrosta
13	8	33.3	8	8	P92384	P92384 hordeum mur
14	8	33.3	8	8	P93973	P93973 eremopyrum
15	8	33.3	8	8	P92215	P92215 amblyopyrum
16	8	33.3	8	8	P93957	P93957 festucopsis
17	8	33.3	8	8	P92222	P92222 bromus iner
18	8	33.3	8	8	P92388	P92388 henrardia p
19	8	33.3	8	8	P92441	P92441 thinopyrum
20	8	33.3	8	8	P92404	P92404 lophopyrum
21	8	33.3	8	8	P93961	P93961 psathyrosta
22	8	33.3	8	8	P93970	P93970 eremopyrum
23	8	33.3	8	8	P93955	P93955 festucopsis
24	8	33.3	8	8	P93965	P93965 secale stri
25	8	33.3	8	8	P92394	P92394 hordeum vul
26	8	33.3	8	8	P92382	P92382 hordeum bra
27	8	33.3	8	8	P93966	P93966 aegilops sp
28	8	33.3	8	8	Q94VA0	Q94va0 varanus sem
29	8	33.3	8	8	P92227	P92227 crithopsis
30	8	33.3	8	8	P92373	P92373 haynaldia v
31	8	33.3	8	8	P92211	P92211 agropyron c
32	8	33.3	8	8	P92428	P92428 peridictyon
33	8	33.3	8	8	P93959	P93959 hordeum ere
34	8	33.3	8	8	P92219	P92219 australopyr
35	8	33.3	8	8	P93985	P93985 aegilops co
36	8	33.3	8	8	P92443	P92443 taeniatheru
37	8	33.3	8	8	P92391	P92391 heteranthel
38	8	33.3	8	8	P93981	P93981 crithodium
39	8	33.3	8	8	P93992	P93992 australopyr
40	8	33.3	8	8	P92426	P92426 pseudoroegn
41	8	33.3	8	8	P92431	P92431 aegilops ta
42	8	33.3	8	8	P92422	P92422 psathyrosta
43	8	33.3	8	13	Q90ZV5	Q90zv5 fulica leuc
44	8	33.3	8	13	P82079	P82079 limnodynast
45	7	29.2	7	10	P82445	P82445 nicotiana t

# ALIGNMENTS

## RESULT 1

Q63480

ID Q63480 PRELIMINARY; PRT; 7 AA.

AC Q63480;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE TR4-NS orphan receptor (Fragment).  
 GN TR4.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96198747; PubMed=8612486;  
 RA Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,  
 RA Detera-Wadleigh S.D.;  
 RT "Splice variants of rat TR4 orphan receptor: differential expression  
 RT of novel sequences in the 5'-untranslated region and C-terminal  
 RT domain.";  
 RL Endocrinology 137:1562-1571(1996).  
 DR EMBL; U59125; AAB02827.1; -.  
 KW Receptor.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 7 AA; 758 MW; 672AA87864005350 CRC64;

Query Match 50.0%; Score 12; DB 11; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GG 2  
 ||  
 Db 4 GG 5

Search completed: October 20, 2003, 17:37:47  
 Job time : 98 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 17:25:34 ; Search time 23 Seconds  
(without alignments)  
8.179 Million cell updates/sec

Title: US-09-807-980-8  
Perfect score: 24  
Sequence: 1 GGGF 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 158

Minimum DB seq length: 0  
Maximum DB seq length: 8

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	18	75.0	8	1	LCK5_LEUMA	P19987, leucophaea	
2	15	62.5	8	1	UF06_MOUSE	P38644 mus musculu	
3	13	54.2	7	1	FAR5_HIRME	P42564 hirudo medi	
4	13	54.2	8	1	AL16_CARMA	P81819 carcinus ma	
5	13	54.2	8	1	ALL9_CARMA	P81812 carcinus ma	
6	12	50.0	4	1	ACH1_ACHFU	P35904 achatina fu	
7	12	50.0	4	1	OCP1_OCTMI	P58648 octopus min	
8	12	50.0	5	1	PAP2_PARMA	P81864 pardachirus	
9	12	50.0	5	1	RE32_LITRU	P82073 litoria rub	
10	12	50.0	6	1	FARP_MONEX	P41966 moniezia ex	
11	12	50.0	8	1	LCK2_LEUMA	P21141 leucophaea	
12	12	50.0	8	1	LCK3_LEUMA	P21142 leucophaea	
13	12	50.0	8	1	ORMY_ORCLI	P82455 orconectes	
14	12	50.0	8	1	VGLG_HSV2B	P81780 herpes simp	
15	12	50.0	8	1	WP1_PERAT	P83195 perkinsus a	
16	11	45.8	8	1	LCK8_LEUMA	P19990 leucophaea	
17	11	45.8	8	1	LMT2_LOCFI	P22396 locusta mig	

18	9	37.5	7	1	FAR1_HELTI	P41871	helisoma tr
19	9	37.5	8	1	ALL5_CYDPO	P82156	cydia pomon
20	8	33.3	8	1	FAR7_ASCSU	P43171	ascaris suu
21	7	29.2	4	1	OCP3_OCTMI	P58649	octopus min
22	7	29.2	8	1	AKHG_GRYBI	P14086	gryllus bim
23	7	29.2	8	1	AKH_TABAT	P14595	tabanus atr
24	7	29.2	8	1	AL15_CARMA	P81818	carcinus ma
25	7	29.2	8	1	AL17_CARMA	P81820	carcinus ma
26	7	29.2	8	1	ALL8_CARMA	P81811	carcinus ma
27	7	29.2	8	1	CCKN_MACEU	P30369	macropus eu
28	7	29.2	8	1	RPCH_PANBO	P08939	pandalus bo
29	6	25.0	3	1	GRWM_HUMAN	P01157	homo sapien
30	6	25.0	4	1	DCML_PSECH	P19916	pseudomonas
31	6	25.0	4	1	EOSI_HUMAN	P02731	homo sapien
32	6	25.0	4	1	FAR3_HIRME	P42562	hirudo medi
33	6	25.0	4	1	FAR4_HIRME	P42563	hirudo medi
34	6	25.0	4	1	FFKA_ANTEL	P58705	anthopleura
35	6	25.0	4	1	FLRF_HIRME	P42561	hirudo medi
36	6	25.0	4	1	FLRN_ANTEL	P58707	anthopleura
37	6	25.0	4	1	FMRF_MACNI	P01162	macrocallis
38	6	25.0	4	1	FYRI_ANTEL	P58706	anthopleura
39	6	25.0	5	1	AL14_CARMA	P81817	carcinus ma
40	6	25.0	5	1	EI03_LITRU	P82099	litoria rub
41	6	25.0	5	1	EI04_LITRU	P82100	litoria rub
42	6	25.0	5	1	FARP_ARTTR	P41853	artioposthi
43	6	25.0	5	1	RE11_LITRU	P82070	litoria rub
44	6	25.0	5	1	RE21_LITRU	P82071	litoria rub
45	6	25.0	5	1	RE31_LITRU	P82072	litoria rub

# ALIGNMENTS

## RESULT 1

LCK5\_LEUMA

ID LCK5\_LEUMA STANDARD; PRT; 8 AA.  
AC P19987;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Leucokinin V (L-V).  
OS Leucophaea maderae (Madeira cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;  
OC Blaberidae; Leucophaea.  
OX NCBI\_TaxID=6988;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Head;  
RX MEDLINE=87052651; PubMed=2877794;  
RA Holman G.M., Cook B.J., Nachman R.J.;  
RT "Isolation, primary structure, and synthesis of leucokinin V and VI:  
myotropic peptides of Leucophaea maderae.";  
RL Comp. Biochem. Physiol. 88C:27-30(1987).  
CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE  
CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).  
CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.

DR PIR; JS0315; JS0315.  
KW Neuropeptide; Amidation.  
FT . MOD\_RES 8 8 AMIDATION.  
SQ SEQUENCE 8 AA; 784 MW; 736365A5B9C865B8 CRC64;

Query Match 75.0%; Score 18; DB 1; Length 8;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGF 4  
| ||  
Db 1 GSGF 4

Search completed: October 20, 2003, 17:36:02  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 17:32:55 ; Search time 39 Seconds  
(without alignments)  
9.863 Million cell updates/sec

Title: US-09-807-980-8  
Perfect score: 24  
Sequence: 1 GGGF 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 604

Minimum DB seq length: 0  
Maximum DB seq length: 8

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%					Description
	No.	Score	Match	Length	ID	
1	18	75.0	5	2	B61445	Leu-enkephalin - b
2	18	75.0	5	2	A61445	Met-enkephalin - b
3	18	75.0	5	2	B61168	cocoonase (EC 3.4.
4	18	75.0	5	2	C53284	T-cell receptor be
5	18	75.0	6	2	A41946	T-cell receptor ga
6	18	75.0	7	2	A60224	Met-enkephalin-Arg
7	18	75.0	7	2	E61491	seed protein ws-5
8	18	75.0	8	2	JS0315	leucokinin V - Mad
9	18	75.0	8	2	PT0509	T-cell receptor be
10	15	62.5	8	2	PC4373	telomeric and tetr
11	12	50.0	3	3	A23751	spinal cord peptid
12	12	50.0	4	2	A32480	achatin-I - giant
13	12	50.0	4	2	A53284	T-cell receptor be



14	12	50.0	4	2	B53284	T-cell receptor be
15	12	50.0	4	2	PT0706	T-cell receptor be
16	12	50.0	4	2	S47552	ubiquitin - rat
17	12	50.0	5	2	C23751	spinal cord peptid
18	12	50.0	5	2	JH0253	gut pentapeptide -
19	12	50.0	5	2	PT0267	Ig heavy chain CRD
20	12	50.0	5	2	D44823	synaptosomal-assoc
21	12	50.0	5	2	PT0669	T-cell receptor be
22	12	50.0	5	2	PT0707	T-cell receptor be
23	12	50.0	5	2	PT0585	T-cell receptor be
24	12	50.0	5	2	PT0717	T-cell receptor be
25	12	50.0	6	2	JU0355	lipopeptide WS1279
26	12	50.0	6	2	PT0280	Ig heavy chain CRD
27	12	50.0	6	2	PT0514	T-cell receptor be
28	12	50.0	6	2	PT0512	T-cell receptor be
29	12	50.0	6	2	PT0605	T-cell receptor be
30	12	50.0	6	2	PT0720	T-cell receptor be
31	12	50.0	6	2	PT0560	T-cell receptor be
32	12	50.0	6	2	PT0723	T-cell receptor be
33	12	50.0	6	2	PT0727	T-cell receptor be
34	12	50.0	6	2	PT0730	T-cell receptor be
35	12	50.0	6	2	A43129	neuropeptide GNFFR
36	12	50.0	7	2	S42407	gramicidin S synth
37	12	50.0	7	2	H33098	180K exoantigen -
38	12	50.0	7	2	PT0523	T-cell receptor be
39	12	50.0	7	2	PT0642	T-cell receptor be
40	12	50.0	7	2	PT0667	T-cell receptor be
41	12	50.0	7	2	PT0666	T-cell receptor be
42	12	50.0	7	2	PT0663	T-cell receptor be
43	12	50.0	7	2	PT0542	T-cell receptor be
44	12	50.0	7	2	PT0683	T-cell receptor be
45	12	50.0	7	2	PT0719	T-cell receptor be

#### ALIGNMENTS

##### RESULT 1

B61445

Leu-enkephalin - blue mussel

C;Species: *Mytilus edulis* (blue mussel)

C;Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 21-Jan-2000

C;Accession: B61445

R;Leung, M.K.; Stefano, G.B.

Proc. Natl. Acad. Sci. U.S.A. 81, 955-958, 1984

A;Title: Isolation and identification of enkephalins in pedal ganglia of *Mytilus edulis* (Mollusca).

A;Reference number: A61445; MUID:84144823; PMID:6583690

A;Accession: B61445

A;Molecule type: protein

A;Residues: 1-5 <LEU>

A;Experimental source: pedal ganglia

C;Keywords: neuropeptide; opioid peptide

Query Match 75.0%; Score 18; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            2 GGF 4  
              | | |  
Db            2 GGF 4

Search completed: October 20, 2003, 17:38:30  
Job time : 39 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 17:37:55 ; Search time 66 Seconds  
(without alignments)  
9.934 Million cell updates/sec

Title: US-09-807-980-8  
Perfect score: 24  
Sequence: 1 GGGF 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 609560 seqs, 163917102 residues

Total number of hits satisfying chosen parameters: 41226

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result	Query					
No.	Score	Match	Length	DB	ID	Description

1	24	100.0	4	11	US-09-866-512A-1	Sequence 1, Appli
2	24	100.0	4	12	US-10-301-499A-1	Sequence 1, Appli
3	24	100.0	4	12	US-10-155-170-8	Sequence 8, Appli
4	24	100.0	5	12	US-10-155-170-5	Sequence 5, Appli
5	24	100.0	7	11	US-09-791-153A-20	Sequence 20, Appl
6	21	87.5	4	12	US-09-800-187-64	Sequence 64, Appl
7	21	87.5	4	12	US-09-800-187-65	Sequence 65, Appl
8	21	87.5	4	12	US-09-800-187-66	Sequence 66, Appl
9	21	87.5	4	12	US-09-800-187-67	Sequence 67, Appl
10	21	87.5	4	12	US-09-800-187-68	Sequence 68, Appl
11	21	87.5	4	12	US-10-301-499A-3	Sequence 3, Appli
12	21	87.5	6	12	US-09-800-187-24	Sequence 24, Appl
13	19	79.2	8	15	US-10-316-534-2	Sequence 2, Appli
14	18	75.0	4	8	US-08-873-601-35	Sequence 35, Appl
15	18	75.0	4	9	US-09-157-748-40	Sequence 40, Appl
16	18	75.0	4	9	US-09-909-652-1	Sequence 1, Appli
17	18	75.0	4	9	US-09-211-691-18	Sequence 18, Appl
18	18	75.0	4	9	US-09-837-969A-42	Sequence 42, Appl
19	18	75.0	4	9	US-09-815-837-106	Sequence 106, App
20	18	75.0	4	9	US-09-815-837-127	Sequence 127, App
21	18	75.0	4	10	US-09-818-247-23	Sequence 23, Appl
22	18	75.0	4	10	US-09-269-439-10	Sequence 10, Appl
23	18	75.0	4	10	US-09-841-321A-42	Sequence 42, Appl
24	18	75.0	4	10	US-09-963-206B-19	Sequence 19, Appl
25	18	75.0	4	10	US-09-753-126-146	Sequence 146, App
26	18	75.0	4	10	US-09-916-940-40	Sequence 40, Appl
27	18	75.0	4	10	US-09-739-068-2	Sequence 2, Appli
28	18	75.0	4	10	US-09-781-988-14	Sequence 14, Appl
29	18	75.0	4	10	US-09-792-630-52	Sequence 52, Appl
30	18	75.0	4	10	US-09-966-976A-19	Sequence 19, Appl
31	18	75.0	4	11	US-09-921-144-30	Sequence 30, Appl
32	18	75.0	4	11	US-09-896-896A-110	Sequence 110, App
33	18	75.0	4	11	US-09-953-351-51	Sequence 51, Appl
34	18	75.0	4	11	US-09-281-495-33	Sequence 33, Appl
35	18	75.0	4	11	US-09-281-495-52	Sequence 52, Appl
36	18	75.0	4	11	US-09-746-371C-4	Sequence 4, Appli
37	18	75.0	4	11	US-09-893-878-14	Sequence 14, Appl
38	18	75.0	4	12	US-10-052-578-142	Sequence 142, App
39	18	75.0	4	12	US-10-084-388A-4	Sequence 4, Appli
40	18	75.0	4	12	US-10-177-725-168	Sequence 168, App
41	18	75.0	4	12	US-10-226-877A-44	Sequence 44, Appl
42	18	75.0	4	12	US-10-262-630-33	Sequence 33, Appl
43	18	75.0	4	12	US-10-133-973-73	Sequence 73, Appl
44	18	75.0	4	12	US-10-301-499A-2	Sequence 2, Appli
45	18	75.0	4	12	US-10-301-499A-4	Sequence 4, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-866-512A-1

; Sequence 1, Application US/09866512A

; Publication No. US20030053954A1

; GENERAL INFORMATION:

; APPLICANT: Meade, Thomas J

; TITLE OF INVENTION: Magnetic Resonance Imaging Agents for the Detection of Physiological

; TITLE OF INVENTION: Agents

; FILE REFERENCE: A-58634-7

; CURRENT APPLICATION NUMBER: US/09/866,512A

; CURRENT FILING DATE: 2001-05-24

; PRIOR APPLICATION NUMBER: US 60/287,619

; PRIOR FILING DATE: 2001-05-26

; PRIOR APPLICATION NUMBER: US 08/460,511

; PRIOR FILING DATE: 1995-06-02

; PRIOR APPLICATION NUMBER: US 08/486,968

; PRIOR FILING DATE: 1995-06-07

; PRIOR APPLICATION NUMBER: US 08/971,855

; PRIOR FILING DATE: 1997-11-17

; PRIOR APPLICATION NUMBER: US 09/134,072

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 09/866,512

; PRIOR FILING DATE: 2001-05-24

; PRIOR APPLICATION NUMBER: US 09/405,046

; PRIOR FILING DATE: 1999-09-27

; PRIOR APPLICATION NUMBER: US 60/063,328

; PRIOR FILING DATE: 1997-10-27

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 4

; TYPE: PRT

; ORGANISM: Unknown

; FEATURE:

; OTHER INFORMATION: could be from any mammal.

US-09-866-512A-1

Query Match 100.0%; Score 24; DB 11; Length 4;

Best Local Similarity 100.0%; Pred. No. 5.5e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGF 4

||||

Db 1 GGGF 4

RESULT 2

US-10-301-499A-1

; Sequence 1, Application US/10301499A

; Publication No. US20030148932A1

; GENERAL INFORMATION:

; APPLICANT: Greene, Mark I.

; APPLICANT: Murali, Mamachandran

; APPLICANT: Berezov, Alan

; TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND METHODS OF USING THE SAME

; FILE REFERENCE: 4040/1K399-US1

; CURRENT APPLICATION NUMBER: US/10/301,499A

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 60/331,935

; PRIOR FILING DATE: 2001-11-21

; NUMBER OF SEQ ID NOS: 152

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1  
; LENGTH: 4  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide  
US-10-301-499A-1

Query Match 100.0%; Score 24; DB 12; Length 4;  
Best Local Similarity 100.0%; Pred. No. 5.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGF 4  
||||  
Db 1 GGGF 4

RESULT 3

US-10-155-170-8

; Sequence 8, Application US/10155170  
; Publication No. US20030171262A1  
; GENERAL INFORMATION:  
; APPLICANT: INOUE, Kazuhiro  
; APPLICANT: SUSAKI, Hiroshi  
; APPLICANT: IKEDA, Masuhiro  
; TITLE OF INVENTION: Drug Complexes  
; FILE REFERENCE: P17306  
; CURRENT APPLICATION NUMBER: US/10/155,170  
; CURRENT FILING DATE: 2002-05-28  
; PRIOR APPLICATION NUMBER: 09/147,342B  
; PRIOR FILING DATE: 1999-03-25  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 4  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Spacer Sequence  
US-10-155-170-8

Query Match 100.0%; Score 24; DB 12; Length 4;  
Best Local Similarity 100.0%; Pred. No. 5.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGF 4  
||||  
Db 1 GGGF 4

Search completed: October 20, 2003, 17:47:29  
Job time : 68 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 17:34:10 ; Search time 29 Seconds  
 (without alignments)  
 5.836 Million cell updates/sec

Title: US-09-807-980-8  
 Perfect score: 24  
 Sequence: 1 GGGF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 61971

Minimum DB seq length: 0  
 Maximum DB seq length: 8

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	24	100.0	4	4	US-09-147-342B-8	Sequence 8, Appli
2	24	100.0	5	1	US-08-257-782-6	Sequence 6, Appli
3	24	100.0	5	1	US-08-577-846-6	Sequence 6, Appli
4	24	100.0	5	2	US-08-079-144-6	Sequence 6, Appli
5	24	100.0	5	4	US-09-147-342B-5	Sequence 5, Appli
6	24	100.0	6	1	US-07-969-307A-11	Sequence 11, Appli
7	24	100.0	6	3	US-08-570-761-2	Sequence 2, Appli
8	24	100.0	7	1	US-07-969-307A-14	Sequence 14, Appli
9	24	100.0	8	1	US-07-969-307A-17	Sequence 17, Appli
10	21	87.5	5	1	US-08-281-195A-26	Sequence 26, Appli
11	21	87.5	5	4	US-09-388-788-1	Sequence 1, Appli

12	21	87.5	7	4	US-09-388-788-2	Sequence 2, Appli
13	19	79.2	5	6	5217869-29	Patent No. 5217869
14	18	75.0	4	1	US-07-664-989B-14	Sequence 14, Appl
15	18	75.0	4	1	US-07-796-243-3	Sequence 3, Appli
16	18	75.0	4	1	US-07-805-727-15	Sequence 15, Appl
17	18	75.0	4	1	US-07-969-307A-4	Sequence 4, Appli
18	18	75.0	4	1	US-08-076-726-4	Sequence 4, Appli
19	18	75.0	4	1	US-08-174-365A-79	Sequence 79, Appl
20	18	75.0	4	1	US-08-174-365A-101	Sequence 101, App
21	18	75.0	4	1	US-08-257-782-36	Sequence 36, Appl
22	18	75.0	4	1	US-08-005-002C-9	Sequence 9, Appli
23	18	75.0	4	1	US-07-972-007-35	Sequence 35, Appl
24	18	75.0	4	1	US-08-351-058A-3	Sequence 3, Appli
25	18	75.0	4	1	US-07-910-867B-6	Sequence 6, Appli
26	18	75.0	4	1	US-07-980-523-3	Sequence 3, Appli
27	18	75.0	4	1	US-08-487-203A-9	Sequence 9, Appli
28	18	75.0	4	1	US-08-346-613-6	Sequence 6, Appli
29	18	75.0	4	1	US-08-676-263-10	Sequence 10, Appl
30	18	75.0	4	1	US-08-488-470A-1	Sequence 1, Appli
31	18	75.0	4	1	US-08-269-929-3	Sequence 3, Appli
32	18	75.0	4	1	US-08-577-846-36	Sequence 36, Appl
33	18	75.0	4	1	US-08-592-029-9	Sequence 9, Appli
34	18	75.0	4	1	US-07-946-239-10	Sequence 10, Appl
35	18	75.0	4	1	US-08-647-618-35	Sequence 35, Appl
36	18	75.0	4	1	US-08-484-635-240	Sequence 240, App
37	18	75.0	4	1	US-08-484-505A-1	Sequence 1, Appli
38	18	75.0	4	2	US-08-453-958-1	Sequence 1, Appli
39	18	75.0	4	2	US-08-484-631-240	Sequence 240, App
40	18	75.0	4	2	US-08-442-461D-35	Sequence 35, Appl
41	18	75.0	4	2	US-08-689-018-1	Sequence 1, Appli
42	18	75.0	4	2	US-08-689-018-3	Sequence 3, Appli
43	18	75.0	4	2	US-08-514-451A-16	Sequence 16, Appl
44	18	75.0	4	2	US-08-249-830-2	Sequence 2, Appli
45	18	75.0	4	2	US-08-670-207-1	Sequence 1, Appli

#### ALIGNMENTS

#### RESULT 1

US-09-147-342B-8

; Sequence 8, Application US/09147342B

; Patent No. 6436912

; GENERAL INFORMATION:

; APPLICANT: INOUE, Kazuhiro

; APPLICANT: SUSAKI, Hiroshi

; APPLICANT: IKEDA, Masuhiro

; TITLE OF INVENTION: Drug Complexes

; FILE REFERENCE: P17306

; CURRENT APPLICATION NUMBER: US/09/147,342B

; CURRENT FILING DATE: 1999-03-25

; PRIOR APPLICATION NUMBER: PCT/JP97/01914

; PRIOR FILING DATE: 1997-06-05

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 8

; LENGTH: 4



; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Spacer Sequence  
US-09-147-342B-8

Query Match 100.0%; Score 24; DB 4; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGF 4  
||||  
Db 1 GGGF 4

RESULT 6

US-07-969-307A-11  
; Sequence 11, Application US/07969307A  
; Patent No. 5463022  
; GENERAL INFORMATION:  
; APPLICANT: INOUE, ET AL.  
; TITLE OF INVENTION: N-ACETYL CARBOXY METHYL CHITOSAN  
; TITLE OF INVENTION: DERIVATIVES AND PROCESS FOR PREPARATION THEREOF  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LARSON AND TAYLOR  
; STREET: 727 SOUTH 23RD STREET  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/969,307A  
; FILING DATE: 16-FEB-1993  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GILLIS, WALTER C.  
; REGISTRATION NUMBER: 22086  
; REFERENCE/DOCKET NUMBER: IV/XX  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-920-7200  
; TELEFAX: 703-892-8428  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
US-07-969-307A-11

Query Match 100.0%; Score 24; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGF 4  
    ||||  
Db 1 GGGF 4

Search completed: October 20, 2003, 17:39:07  
Job time : 31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 17:17:59 ; Search time 42 Seconds  
(without alignments)  
15.117 Million cell updates/sec

Title: US-09-807-980-8  
Perfect score: 24  
Sequence: 1 GGGF 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	%		Query Match Length	DB	ID	Description
	Score					
1	24	100.0	4	20	AAAY04403	Spacer peptide #6.
2	24	100.0	4	21	AAAY91119	Drug delivery syst
3	24	100.0	4	21	AAAY49437	Spacer peptide mol
4	24	100.0	4	23	ABG32178	Drug complex assoc
5	24	100.0	4	23	ABB06035	Spacer peptide #8.
6	24	100.0	4	24	ABP55660	Target substance/p
7	24	100.0	5	20	AAAY04402	Spacer peptide #5.
8	24	100.0	5	21	AAB03528	Drug delivery syst
9	24	100.0	5	22	AAU03577	ICC13 pentapeptide
10	24	100.0	5	23	ABG32175	Drug complex assoc
11	24	100.0	5	23	ABB06032	Spacer peptide #5.
12	24	100.0	6	16	AAR64944	Peralkylated oligo
13	24	100.0	6	16	AAR64945	Peralkylated oligo
14	24	100.0	7	22	AAU08364	Antibody heavy cha
15	24	100.0	7	22	AAB82893	Anti-human CD154 a
16	24	100.0	8	19	AAW54524	High affinity Kb b
17	24	100.0	9	16	AAR72894	E. coli PPIase-bet
18	24	100.0	9	16	AAR72938	E. coli PPIase-bet
19	24	100.0	9	17	AAR99268	Leukocyte cell-der
20	24	100.0	9	23	AAU95037	Human novel protei
21	24	100.0	9	24	ABR17409	Human cancer-relat
22	24	100.0	9	24	ABR17442	Human cancer-relat
23	24	100.0	9	24	ABR17463	Human cancer-relat
24	24	100.0	9	24	ABR17597	Human cancer-relat
25	24	100.0	9	24	ABR17822	Human cancer-relat
26	24	100.0	9	24	ABR17994	Human cancer-relat
27	24	100.0	9	24	ABR18045	Human cancer-relat
28	24	100.0	9	24	ABR18171	Human cancer-relat
29	24	100.0	9	24	ABR18195	Human cancer-relat
30	24	100.0	9	24	ABR18210	Human cancer-relat
31	24	100.0	9	24	ABR18393	Human cancer-relat
32	24	100.0	9	24	ABR18601	Human cancer-relat
33	24	100.0	9	24	ABR18620	Human cancer-relat
34	24	100.0	9	24	ABR18622	Human cancer-relat
35	24	100.0	10	16	AAR72949	E. coli PPIase-alp
36	24	100.0	10	16	AAR72905	E. coli PPIase-alp
37	24	100.0	10	22	AAG94804	Human complementar
38	24	100.0	10	22	AAG94806	Human complementar
39	24	100.0	10	22	AAG94808	Human complementar
40	24	100.0	10	22	AAG97296	Human complementar
41	24	100.0	10	22	AAG97298	Human complementar
42	24	100.0	10	22	AAG97300	Human complementar
43	24	100.0	10	22	AAG83808	Arabidopsis thalia
44	24	100.0	10	22	AAG83983	Arabidopsis thalia
45	24	100.0	10	22	AAG84059	Arabidopsis thalia

# ALIGNMENTS

RESULT 1

AAY04403

ID AAY04403 standard; peptide; 4 AA.

XX

AC AAY04403;

XX

DT 25-JUN-1999 (first entry)

XX

DE Spacer peptide #6.

XX

KW Spacer; medicine complex; carboxyl-4C alkyl pullulan polyalcohol;  
KW tumour.

XX

OS Synthetic.

XX

PN JP11092405-A.

XX

PD 06-APR-1999.

XX

PF 19-SEP-1997; 97JP-0254780.

XX

PR 19-SEP-1997; 97JP-0254780.

XX

PA (DAUC ) DAIICHI PHARM CO LTD.

PA (DDSK-) DDS KENKYUSHO KK.

XX

DR WPI; 1999-283504/24.

XX

PT New medicine complex comprising pullulan alcohol - useful for its  
PT tumour site selectivity

XX

PS Disclosure; Page 6; 12pp; Japanese.

XX

CC The present invention describes a medicine complex comprising a  
CC carboxyl-4C alkyl pullulan polyalcohol combined with the residue of a  
CC medical compound through a spacer consisting of an amino acid or a  
CC spacer especially of 2 to 8 amino acids peptide-bonded. Also described  
CC are: (1) a carrier for medicine delivery for combining a medical  
CC compound consisting of a carboxyl-4C alkyl pullulan polyalcohol; and (2)  
CC use of a carboxyl-4C alkyl pullulan polyalcohol for the preparation of a  
CC medicine complex containing a carboxyl-4C alkyl pullulan polyalcohol  
CC combined to the residue of a medical compound optionally through a  
CC spacer. The medicine complex is useful for its tumour site selectivity.  
CC The present sequence represents an example of a spacer peptide given in  
CC the present invention.

XX

SQ Sequence 4 AA;

Query Match 100.0%; Score 24; DB 20; Length 4;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGF 4

||||

Db 1 GGGF 4

RESULT 2

AA91119

ID AA91119 standard; peptide; 4 AA.

XX

AC AA91119;

XX

DT 06-OCT-2000 (first entry)

XX

DE Drug delivery system compound N-terminal tetrapeptide #2.

XX

KW Drug delivery system; DDS; saccharide; peptidase; cytostatic;

KW antiinflammatory; modified carboxy 1-4C alkyl dextran polyalcohol;

KW hydrolysate; antitumour; liver cancer.

XX

OS Unidentified.

XX

PN WO200025825-A1.

XX

PD 11-MAY-2000.

XX

PF 29-OCT-1999; 99WO-JP06016.

XX

PR 30-OCT-1998; 98JP-0310130.

PR 19-NOV-1998; 98JP-0329272.

XX

PA (DAUC ) DAIICHI PHARM CO LTD.

XX

PI Susaki H, Inoue K, Kuga H, Ikeda M, Shiose Y, Korenaga H;

XX

DR WPI; 2000-365409/31.

XX

PT New drug delivery system compounds comprise saccharide compound

PT modified carboxy alkyl dextran polyalcohol bonded to antitumor or

PT antiinflammatory agent -

XX

PS Claim 33; Page 52; 64pp; Japanese.

XX

CC The present invention describes a drug delivery system (DDS) compound

CC comprising a saccharide compound modified carboxy 1-4C alkyl dextran

CC polyalcohol bonded to a drug compound. Also described is a method of

CC assaying DDS compounds with a drug bonded to a polymer carrier via a

CC spacer containing 2-8 amino acids, comprising assaying a hydrolysate

CC obtained by treating the DDS compound with a peptidase. The compound

CC is used as a drug delivery system for administering e.g. antitumour

CC and antiinflammatory drugs, especially for treating liver cancer. The

CC assay can be used to monitor distribution and blood levels of the drug

CC to allow accurate dosing. The carrier increases bioavailability of drug

CC and allows the drug levels in the body to be readily assayed. The

CC present sequence represents a specifically claimed N-terminal

CC tetrapeptide which is used in the exemplification of the present

CC invention.

XX

SQ Sequence 4 AA;

Query Match 100.0%; Score 24; DB 21; Length 4;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGF 4  
    ||||  
Db 1 GGGF 4

RESULT 3

AA49437

ID AA49437 standard; peptide; 4 AA.

XX

AC AA49437;

XX

DT 17-MAR-2000 (first entry)

XX

DE Spacer peptide molecule in a drug composite.

XX

KW Drug composite; spacer; drug delivery system; antitumor;  
KW antiinflammatory.

XX

OS Synthetic.

XX

PN WO9961061-A1.

XX

PD 02-DEC-1999.

XX

PF 21-MAY-1999; 99WO-JP02681.

XX

PR 22-MAY-1998; 98JP-0140915.

XX

PA (DAUC ) DAIICHI PHARM CO LTD.

XX

PI Susaki H, Inoue K, Kuga H;

XX

DR WPI; 2000-072550/06.

XX

PT Drug composite comprises carrier bound to drug via spacer useful as  
PT drug delivery systems -

XX

PS Claim 17; Page 41; 53pp; Japanese.

XX

CC The invention provides drug composites comprising a polymer carrier  
CC bound through an amino acid spacer to a drug. The drug composite  
CC comprises a compound of formula A-R-NH-Y-CH<sub>2</sub>-O-CO-Q where, A = polymer  
CC carrier for a drug; R = spacer comprising 1-8 amino acid molecules bound  
CC to each other through a peptide linkage; Y = optionally substituted  
CC phenylene; Q = residue of drug. The drug composites are used as drug  
CC delivery systems for antitumor or antiinflammatory agents. The composites  
CC give rapid and regioselective release of drug thus increasing activity  
CC and maximum tolerated dose of drug. The present sequence represents a  
CC spacer peptide of the drug composite.

XX

SQ Sequence 4 AA;

Query Match 100.0%; Score 24; DB 21; Length 4;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGF 4  
||||  
Db 1 GGGF 4

RESULT 4

ABG32178

ID ABG32178 standard; Peptide; 4 AA.

XX

AC ABG32178;

XX

DT 29-NOV-2002 (first entry)

XX

DE Drug complex associated peptide spacer #8.

XX

KW Drug complex; spacer; carboxy(1-4C)alkyldextran polyalcohol; cancer;  
KW inflammation; antineoplastic; antiinflammatory.

XX

OS Synthetic.

XX

PN US6436912-B1.

XX

PD 20-AUG-2002.

XX

PF 25-MAR-1999; 99US-0147342.

XX

PR 06-JUN-1996; 96JP-0144421.

PR 05-JUN-1997; 97WO-JP01914.

XX

PA (DAUC ) DAIICHI PHARM CO LTD.

PA (DRUG-) DRUG DELIVERY SYSTEM INST LTD.

XX

PI Inoue K, Susaki H, Ikeda M, Kuga H, Kumazawa E, Togo A;

XX

DR WPI; 2002-673238/72.

XX

PT A drug complex used as antineoplastic or antiinflammatory, comprises a  
PT carboxy(1-4C)alkyldextran polyalcohol residue bound to a drug compound  
PT via an amino acid or 2-8 amino acid peptide spacer -

XX

PS Example 2; Column 15; 41pp; English.

XX

CC The invention describes a drug complex comprising a  
CC carboxy(1-4C)alkyldextran polyalcohol and a residue of a drug compound  
CC bound to each other with a spacer comprising an amino acid or  
CC peptide-bonded 2-8 amino acids. The drug complex is used as a medicament  
CC for the treatment of various cancers or inflammation. The drug complex is  
CC capable of site-selectively delivering the antineoplastic or  
CC antiinflammatory to a tumorous site or inflammatory site. The drug can be  
CC retained in the blood for a long period of time and has reduced toxicity.  
CC Since protease is expressed at tumorous and inflammatory sites, the  
CC spacer used here is readily hydrolyzed to allow a released drug compound  
CC to exhibit its efficiency. This sequence represents a peptide spacer used  
CC in the creation of a drug complex for treatment of inflammation and  
CC cancer.

XX

SQ Sequence 4 AA;



Query Match 100.0%; Score 24; DB 23; Length 4;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGF 4  
    ||||  
Db 1 GGGF 4

RESULT 12

AAR64944

ID AAR64944 standard; peptide; 6 AA.

XX

AC AAR64944;

XX

DT 25-MAR-2003 (updated)

DT 12-SEP-1995 (first entry)

XX

DE Peralkylated oligopeptide.

XX

KW Peralkylated oligopeptides; binding assays; hydrolysis resistance;  
KW antimicrobial drug design.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Modified-site 1

FT /note= "trimethylammonium N-terminal nitrogen"

FT Modified-site 6

FT /note= "C-terminal carboxamide group"

FT Modified-site 1..6

FT /note= "A methyl group on each amido nitrogen of  
FT the backbone and appropriately methylated  
FT side chains, at total of 8 added methyls."

XX

PN WO9500539-A1.

XX

PD 05-JAN-1995.

XX

PF 10-JUN-1994; 94WO-US06554.

XX

PR 17-JUN-1993; 93US-0079144.

PR 09-JUN-1994; 94US-0257782.

XX

PA (HOUG-) HOUGHTEN PHARM INC.

XX

PI Blondelle S, Houghten R, Ostresh JM, Houghten RA;

XX

DR WPI; 1995-052001/07.

XX

PT Sets and libraries of per:alkylated oligopeptide derivs. -  
PT including new cpds., with selected amino acids at partic.  
PT positions, used in binding assays to identify sequence with  
PT optimum binding to e.g. cellular receptors.

XX

PS Example 4; Page 105; 162pp; English.

XX

CC AAR64940-R64959 are peralkylated oligopeptides (POs), which  
CC together comprise a PO library. PO libraries can be used to  
CC assay peptides for preferential binding to an acceptor, esp. for  
CC drug (specifically antimicrobial) design. The POs are resistant  
CC to enzymatic hydrolysis, also POs that are toxic to gram positive  
CC bacteria do not cause haemolysis of human blood.  
CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 6 AA;

Query Match 100.0%; Score 24; DB 16; Length 6;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGF 4  
    ||||  
Db 2 GGGF 5

Search completed: October 20, 2003, 17:23:31  
Job time : 44 secs